Open Praxis, vol. 6 issue 2, April–June 2014, pp. 189–197 (ISSN 2304-070X)
2014 OCWC Global Conference Selected Papers

A Framework to Integrate Public, Dynamic Metrics Into an OER Platform

Jaclyn Zetta Cohen, Kathleen Ludewig Omollo & Dave Malicke
University of Michigan (United States of America)
jczetta@umich.edu, kludewig@umich.edu & dmalicke@umich.edu

Abstract

The usage metrics for open educational resources (OER) are often either hidden behind an authentication system or shared intermittently in static, aggregated format at the repository level. This paper discusses the first year of University of Michigan's project to share its OER usage data dynamically, publicly, to synthesize it across different levels within the repository hierarchies, and to aggregate in a method inclusive of content hosted on third-party platforms. The authors analyze their user research with a target audience of faculty authors, multimedia specialists, librarians, and communications specialists. Next, they explore a stratified technical design that allows the dynamic sharing of metrics down to the level of individual resources. The authors conclude that this framework enables sustainable feedback to OER creators, helps to build positive relationships with creators of OER, and allows the institution to move toward sharing OER on a larger scale.

Keywords: analytics; APIs; CMS; data; metric; OER; user-research

Introduction

Open.Michigan was launched in April 2008 as an umbrella open education initiative at the University of Michigan (U-M). Open.Michigan has two primary goals: to sustain a thriving culture of sharing knowledge at U-M, and to provide comprehensive public access to all of U-M's scholarly output. The Open.Michigan website (http://open.umich.edu/) serves as the primary portal to Open Educational Resources (OER) produced by U-M. It acts as both a repository and a referratory, with some content hosted within the system and some links to content hosted on third-party platforms (e.g. YouTube, Slideshare, Internet Archive). The Open.Michigan website contains OER and learning experiences from eighteen U-M schools and colleges, as well as materials from several partner institutions. The collection includes thousands of learning materials, including over 100 courses, dozens of learning modules, over 30 books, over 1500 videos, and over 150 translated caption tracks for videos. The most comprehensive departmental collections are the Medical School, the School of Information, the School of Dentistry, and the Library.

As an initiative that promotes a culture of sharing knowledge, we believe in following fully open practices for our operations. This includes incorporating principles of public access and open licenses into our learning materials (http://open.umich.edu/education), software (https://github.com/openmichigan/), scholarly publications (http://open.umich.edu/about/infokit/), and process documentation (http://open.umich.edu/share/).

Over the past several years, we have seen an increase in the number of requests from authors and departmental administrators for customized usage reports for their learning materials in the Open.Michigan collection. Since 2008, Google Analytics has been the primary method that we use to monitor the Open.Michigan website traffic. This means that there is extensive historical usage data for Open.Michigan resources, but that data has been traditionally restricted to those with access to our Google Analytics account (Rodgers, 2011). When usage details were requested, a member

Reception date: 28 February 2014 • Acceptance date: 8 April 2014
DOI: http://dx.doi.org/10.5944/openpraxis.6.2.118

of our staff usually collected it through the Google Analytics web interface. This approach was not scalable due to the level of individual effort required for each request.

We needed a way to more systematically communicate the usage of our OER. In 2011, we created a dashboard using Geckoboard (http://www.geckoboard.com/), primarily for internal use, which provided a single-screen view of aggregate usage data. In early 2012, we wrote python scripts to collect usage data for a subset of our resources using the APIs for Google Analytics and YouTube to prepare spreadsheets to be analyzed by volunteers during a weekend-long data analysis event (http://open.umich.edu/education/si/resources/datadive/winter2012).

In 2013, we launched a project to extend those principles to increase access to the usage data for our collections (http://openmi.ch/dyn-metrics). Making our usage data publicly available would provide authors, departmental administrators, and the general public with access to that information for use as they deemed fit. The goal of our "Dynamic Metrics and Analytics Project" is to build a framework for sharing usage data in more useful, accessible, and timely ways for our audience. In early 2013, we released a button on the right margin of all course and resource pages. When the button is clicked, it reveals aggregate, lifetime metrics (updated daily) for that course or resource. We proceeded with user interviews to prioritize subsequent features and designed the scaffolding to integrate new functionality into the framework.

In this paper we discuss the past year of this project, including our user research and technical design. We explore our ability to share dynamic metrics down to the level of individual courses and resources. We conclude that this stratified framework helps us to deliver sustainable feedback to OER creators on their efforts toward openly sharing materials, helps us to build positive relationships with creators of OER, and allows us to move toward sharing OER on a larger scale.

User Research

We conducted semi-structured interviews in order to identify what metrics and indicators were of particular interest to our audience and why. Twenty interviewees were selected from a pool of faculty, students, and staff who had recently worked with Open.Michigan to create, support, or promote OER in some way. Of the people contacted, nine individuals responded to our requests for interviews: two librarians, two faculty members, three multimedia specialists, and two communication specialists. Eight of those people were from the University of Michigan, and one was from an Open.Michigan partner organization for the African Health OER Network (http://oerafrica.org/healthoer/). The semi-structured interviews focused on usage data and the production of OER. All the interviews aimed to identify what, if any, metrics were of particular interest to our audience, and why.

Each interview was 30–60 minutes. We used a loosely structured interview protocol to guide the direction of these interviews, and allowed discussions to evolve, if relevant, within the interview time. Most were conducted at the Open.Michigan office, two were conducted in the respective faculty or staff member's office elsewhere on the University of Michigan campus, and one with a colleague outside the university was conducted as an audio conference via Skype.

Depending on interview length, some interviews included a card-sorting exercise with 15 cards (http://openmi.ch/dyn-metrics-cards) that represented potential features for a metrics display tool. Each interviewee was asked to rank the potential feature cards in their personal order of importance. Some of the features discussed included: "a time series chart of course page views over the past thirty days," "a comparison of metrics over the past month and the past year," "total aggregate numbers of views and downloads," and "a summary in PDF format regularly emailed to me.

The card-sorting exercise allowed us to examine users' priorities in more detail (Holtzblatt & Beyer, 2013). While in a more free-form user interview, an interviewee might express an interest in information communication via any medium, or across many different timelines (30 days, 1 year, full history), the requirement of prioritizing these features in the card-sorting exercise was helpful to give context to the interview data. For instance, a PDF presentation of metrics was repeatedly brought up in comparison to the Google Analytics PDFs that one support staff member regularly sent to her department. Through the card-sorting exercise we surmised that an online view of charts and aggregate numbers would be a better fit for her use cases overall.

The user interviews and card sorting exercises determined our priorities for subsequent metrics features. For instance, access to metrics on more third-party platforms such as Slideshare and the ability to look at different time ranges on time-series charts are both features approaching release at the time of this writing. Additionally, the user interviews and the card sorting exercises surfaced two primary motivations for public, dynamic OER usage metrics, which are explored in the next section.

Analysis of User Interviews

An analysis of the user interview notes revealed two primary motivations for sharing and accessing dynamic usage metrics for OER: to strengthen relationships with authors and to justify the effort and expense of OER practices.

Readily Available Metrics Strengthen Relationships with OER Authors

The librarians, multimedia specialists, and communication specialists with whom we spoke suggested that publicly displayed metrics could be used to initiate conversations with authors. A conversation with supporting metrics can demonstrate the tangible impact of published OER, thereby creating "less disconnect from [the] content" and its author(s). One interviewee added, "having numbers available in passing such as on the landing page is good way to demonstrate that their OER has value." One interviewee noted that the metrics data and other small gestures such as highlighting learning materials as featured resources are "bits of 'yay' and kudos for creators, it gives people that encouragement." A communications specialist who regularly prepares handouts about OER to share at campus events for prospective and current faculty, researchers, and students added, "[It] can't hurt every now and then to say what's super popular."

The online audience for the material may not be visible to the author once an OER is published. This is especially true for OER that are published on platforms that are primarily for use by individual self-paced learners. Publicly displaying metrics for published OER in a centralized location, such as the Open.Michigan website, provides a space for the authors to view the frequency with which the public accessed and downloaded their materials. One librarian interviewed commented that revealing these stats demonstrates that even content that was shared several years ago may still be actively used. One multimedia specialist noted that publicly displaying metrics "provides a structure where people can interact with it in more ways."

These interviewees suggested that strengthening the relationship between the authors and their OER may also strengthen the authors' relationship with the campus OER community and the Open.Michigan staff in particular. Furthermore, the interviewees believe that strengthening relationships between the authors, their published OER, and Open.Michigan may someday encourage future authorship of OER. One interviewee noted that such sharing of metrics could "document impact and may persuade people to join the open content party."

Readily Available Metrics Are Evidence to Support Effort and Expense for OER

Developing learning materials involves time and money. Interviewees reported that metrics can be important evidence to justify the effort and expense for those activities—both on a personal level for the authors and on an organizational level for allocating resources.

Faculty we interviewed stated that it was satisfying to know if their OER are being used and were particularly interested in the geographic makeup of their audience. One faculty member printed and posted a world map of views in various countries for one of his open courses, which he has kept on his office door for over a year.

Both faculty members also expressed the importance of granularity of metrics—at the level of a whole course and at the level of individual files. Having that granularity was "personally satisfying" and "very important"; stratified level detail can indicate the level of audience engagement. For example, one author wanted to know if there were particular materials within a course that were especially popular and whether the audience was accessing materials beyond a cursory view of the syllabus. One of the librarians echoed that sentiment, and added that having that level of granularity lends itself well to benchmarking against peers: "it would be great to be able to say it was a bestseller" for a particular topic. One librarian echoed that simply "satisfying curiosity is good."

Interviewees who worked in communications or multimedia support expressed the importance of having readily accessible metrics to report value back to the organization for its OER collections. One librarian added, "how can this data be used to make decisions about how we invest in things?" One multimedia specialist expressed he has to "be ready to make our case with numbers" when explaining the motivations and results of open licenses. Another added, "as soon as you add a number, people see it as more tangible."

Being able to collect and stratify metrics from OER hosted across multiple hosting platforms can also be used to inform decisions about which platforms to prioritize for distribution: "We use the stats to communicate to management where to spend money on this stuff. Maybe a course is more popular on iTunes U than YouTube. That helps us decide where to put the subsequent course. Is it worth the staff time to put in both places?"

Technical Design

The technical architecture used by an organization to host and reference OER is tied closely to the organization's ability to share detailed usage data for its OER. The Open.Michigan website was built using a customized distribution of the open source content management system (CMS) Drupal (http://www.drupal.org/). Drupal's hierarchical structures allow us to provide metrics at varying levels of that hierarchy. Drupal relies upon a relational database (MySQL, in our case). In our Drupal customization, called OERbit (http://open.umich.edu/oerbit), the database uses a tree-node structure. This structure allows us to identify metrics for individual courses and resources and for aggregating usage data on content associated specifically with that course or resource or that unit.

Each node excepting the single site root has a parent node, most nodes have peer nodes, and individual materials act as leaf nodes. For example, the course M2 Cardio (http://open.umich.edu/education/med/m2/cardio) is a child of the second-year medical school curriculum and a grandchild of the Medical School Unit. The M2 cardio course contains several dozen materials, such as a syllabus, a schedule, lecture slides, lecture videos, and lecture notes. It also has peer courses in the M2 curriculum. This structure is shown in Figure 1.

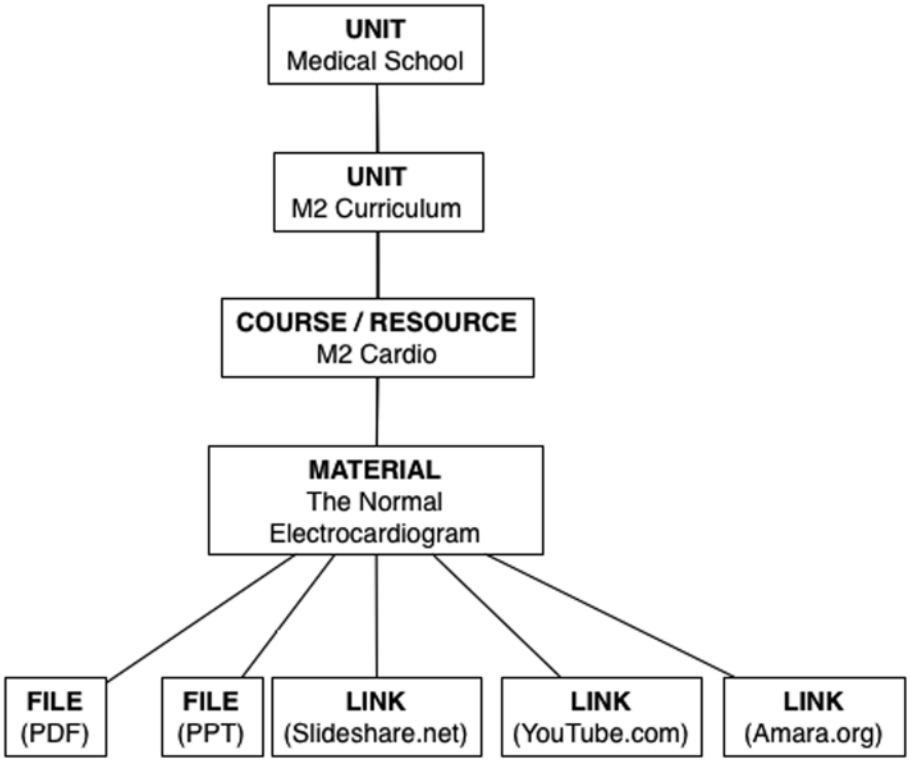

Figure 1: Excerpt from Tree-Node Structure of Units, Courses, and Materials in Open.Michigan OER Platform.

Note: An academic unit may have sub-units. Each unit may have courses or resources, with courses being more a formally structured learning experience. Each course or resource has materials, which may be stored as files in the database or included as external links to OER hosted on a number of external websites.

Courses and resources on the Open.Michigan website may also include links to OER hosted on a number of external websites (e.g. YouTube, SlideShare). The tree-node structure is also useful for collecting metrics from these third-party platforms. Third-party content is included as materials with links to external websites. By parsing the hyperlink, we can determine both the service and the ID for a particular resource. For example, the M2 Cardio course includes a material that is a lecture video for The Normal Electrocardiogram. The hyperlink is http://www.youtube.com/watch?v=BuH_5Wvcd8k. The service is YouTube and the video ID is BuH_5Wvcd8k. Through YouTube's Application Programmable Interfaces (API), we use the video ID to collect usage details for that particular video. Those details are aggregated with the other YouTube videos for the course. Figure 2 shows the aforementioned button, which is displayed by default. It may be hidden by clicking the X in the corner. Figure 3 shows the dashboard currently under development, both for the M2 Cardio course. Currently, YouTube is the only third-party API integrated into the Open.Michigan metrics display. Integration with the APIs from Slideshare.net and Amara.org are currently under development.

By leveraging REST APIs, we can collect these metrics from third-party services and use Drupal modules to drive cron jobs that update metrics information daily. This process is run most efficiently during a time when the website receives the least amount of traffic.

By integrating the metrics from these third-party platforms, Open.Michigan is able to provide a more comprehensive snapshot of the metrics for the OER associated with a particular course or resource. This structure provides course or resource level metrics for individual content creators

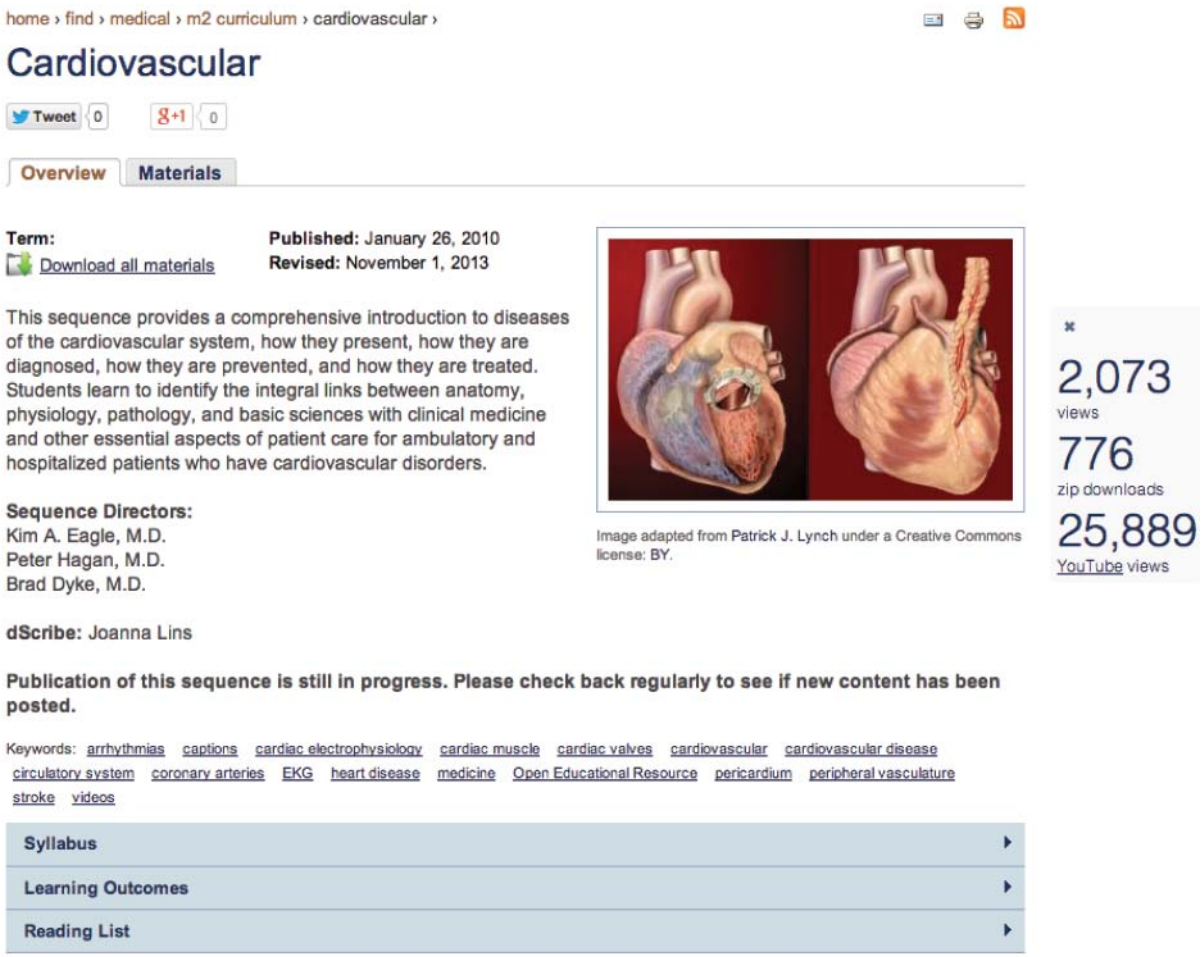

Figure 2: Dynamic Metrics Button Publicly Displayed on Course Landing Page for M2 Cardio.

Note: Upon click, users may show or hide the dynamic metrics button, which contains views and zip downloads from Google Analytics and sum views for all YouTube videos in the course.

and staff, who generally create one or a few particular courses or resources and are curious about those in particular rather than the entire collection or the unit to which those courses or resources belong.

Many OER providers share site-wide aggregate usage data summarized by month or by year, but usually this information is shared in a static .PDF format (OpenCourseWare Consortium 2014; MIT OpenCourseWare 2013). Our metrics framework allows Open.Michigan to share metrics in a more detailed manner, as discussed, and much more dynamically.

Based on the user interviews, we are currently integrating the metrics APIs from SlideShare and Amara into the dashboards. In addition, we are developing export and API functionality as additional methods to share the aggregated usage metrics for courses and for academic units.

Metrics Enable Detection of Patterns and Deeper Analysis of Use

Our user interviews point us toward more specific future directions for our projects, both conceptually and technologically.

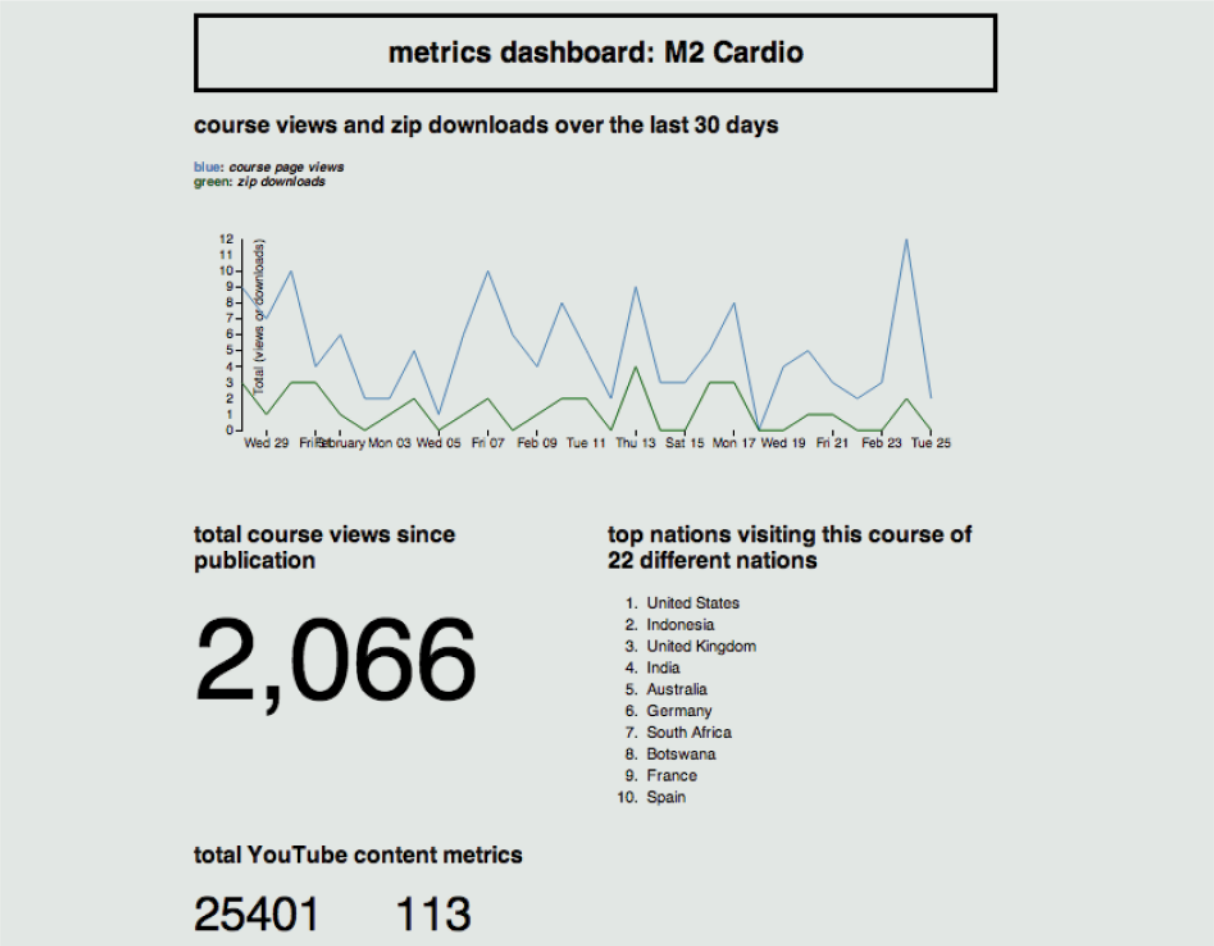

Figure 3: Dynamic Metrics Dashboard In Development for Course Landing for M2 Cardio.

Note: From Google Analytics, the dashboard shows views and zip downloads over the past 30 days, total lifetime views, total number of nations, and the names of the top 10 nations. From YouTube, the dashboard shows total views and number of comments.

Metrics can inform quantitative investigation, but they do not answer "why" questions (e.g. "why does this course have so many downloads compared to this other?") Analytics, which we define as meaningful patterns derived from metrics, are the answers to the whys. Metrics are part of the pathway to analytics and understanding audience use and engagement (Muramatsu & Caswell 2010; Alevizou 2012). Synthesis and analysis of social context is necessary to progress from metrics to analytics in order to determine answers to "why" and "who" questions, about which many of our interviewees were curious.

People wanted to know the whys and the hows, while metrics can tell only the counts and the whats. We know that efficient gathering of numbers, and time-stamped data, is the first step toward diving deeper into whys and hows of OER use. Are these OER helping people? A faculty member noted that if people were using things he created, it would be "good to know where, it would be good to know why."

One faculty member noted that usage data is particularly interesting "not as a cumulative [measure], but as a pattern over time." Looking at when school terms begin and end (both locally

and in other countries with frequent access) and measuring rises or dips in access of certain materials coinciding with those times. Faculty and librarians both expressed interest in comparing similar courses and resources and discovering which courses and resources were seem to have similar frequencies of use.

Many interviewees noted future possibilities given an available supply of usage metrics, which allowed us to *do something.* Many of these future possibilities hinged on extracting interesting patterns from data over time. A multimedia specialist noted that it would be nice to be able to give anyone who wanted them the tools for a spontaneous data analysis event like the one previously mentioned in 2012. Like the interviewee comments referred to earlier about how this data can be used to make decisions about how we invest in things, questions about the raw metrics we are gathering can lead us toward making decisions about what to build next such that it will provide value to those in our community.

Conclusion

Metric-sharing is currency in relationships between OER-publishing-platforms and the faculty, staff, and students of universities and businesses who create OER. This is not the only benefit, but it has been a clear one for us: we are giving a population we have identified as a primary audience something they want. Our experience and user research confirm this.

The technical architecture an organization uses to host and reference OER is tied closely to the organization's ability to share detailed usage data for its OER. Our use of the hierarchical structure of a Drupal-based platform allows us to easily provide metrics for individual courses or resources. Metrics on a small scale are interesting, especially to their creators. Metrics can inform quantitative investigation, but they do not answer "why" questions (e.g. "why does this course have so many downloads compared to this other?"). To find analytics, metrics are necessary, but not sufficient. Due to the structure by which OERbit platform stores metadata, Open.Michigan can group OER and its associated metrics in various ways. This positions Open.Michigan to progress from dynamic metrics to dynamic analytics in the future.

Much like seeing nodding and note-taking when you speak in front of an auditorium, seeing evidence of views, downloads, or comments where your OER are published is validation, evidence that there is some likelihood your effort provides real value to others. Having that evidence allows many to justify the additional effort it may take to openly license educational materials to supervisors and administrators. Open.Michigan's sharing of individual OER metrics sustains development of open resources and allows an open education initiative, such as ours, to build strong relationships with its surrounding community and thus support the development and sharing of OER on a larger scale.

Acknowledgements

This paper was presented at the 2014 OpenCourseWare Consortium Global Conference, held in Ljubljana (Slovenia) in April 23th-25th 2014 (http://conference.ocwconsortium.org/2014), with whom *Open Praxis* established a partnership. After a pre-selection by the Conference Programme Committee, the paper underwent the usual peer-review process in *Open Praxis*.

The authors wish to thank the following individuals for their support and guidance throughout the dynamic metrics and analytics project: Pieter Kleymeer, Kevin Coffman, Bob Riddle, Susan Topol, Stephanie Dascola, Margaret Ann Murphy, Trisha Paul, Emily Puckett Rodgers, Pierre Clement, and Michael Hess.

References

Alevizou, P. (2012). Open to interpretation?: productive frameworks for understanding audience engagement with OER. *Cambridge 2012: Innovation and Impact—Openly Collaborating to Enhance Education, a joint meeting of OER12 and OpenCourseWare Consortium Global 2012*, 16–18 April 2012, Cambridge, UK. Retrieved from http://oro.open.ac.uk/id/eprint/33452

Holtzblatt, K. & Beyer, H. R. (2013). Contextual Design. In M. Soegaard & R. F. Dam (eds.). *The Encyclopedia of Human-Computer Interaction* (2nd Ed). Aarhus, Denmark: The Interaction Design Foundation. Retrieved from http://www.interaction-design.org/encyclopedia/contextual_design.html

Massachusetts Institute of Technology, MIT OpenCourseWare. (2013). *Technology: Leveraging Existing Course/Content Management Systems*. Retrieved from http://www.ocw.nur.ac.rw/OcwWeb/HowTo/tech-leverage-cms.htm

Muramatsu, B. & Caswell, T. (2010). Plagiarism is Good: Moving from Access to Use as Metrics for OCW/OER Use and Reuse. *Open Education 2010 Conference*. Barcelona, Spain. (Vol. 3, November 2010). Retrieved from http://www.slideshare.net/bmuramatsu/plagiarism-is-good-moving-from-access-to-use-as-metrics-for-ocwoer-use-and-reuse

OpenCourseWare Consortium. (2014). *OpenCourseWare Consortium: Evaluation and Metrics*. Retrieved from http://www.ocwconsortium.org/resources/toolkits/evaluation

Rodgers, E. (2011). Measuring Our Impact: the Open.Michigan Initiative. In *Proceedings of OpenCourseWare Consortium Global 2011: Celebrating 10 Years of OpenCourseWare*. Retrieved from http://hdl.handle.net/2027.42/91308